

# Fundamental Limits to Position Determination by Concentration Gradients

Filipe Tostevin<sup>1</sup>, Pieter Rein ten Wolde<sup>2</sup> and Martin Howard<sup>1\*</sup>

<sup>1</sup> Department of Mathematics, Imperial College London,  
South Kensington Campus, London SW7 2AZ, UK

<sup>2</sup> FOM Institute for Atomic and Molecular Physics (AMOLF),  
Kruislaan 407, 1098 SJ, Amsterdam, The Netherlands

\* To whom correspondence should be addressed.  
E-mail:martin.howard@imperial.ac.uk

February 1, 2008

## Abstract

Position determination in biological systems is often achieved through protein concentration gradients. Measuring the local concentration of such a protein with a spatially-varying distribution allows the measurement of position within the system. In order for these systems to work effectively, position determination must be robust to noise. Here, we calculate fundamental limits to the precision of position determination by concentration gradients due to unavoidable biochemical noise perturbing the gradients. We focus on gradient proteins with first order reaction kinetics. Systems of this type have been experimentally characterised in both developmental and cell biology settings. For a single gradient we show that, through time-averaging, great precision can potentially be achieved even with very low protein copy numbers. As a second example, we investigate the ability of a system with oppositely directed gradients to find its centre. With this mechanism, positional precision close to the centre improves more slowly with increasing averaging time, and so longer averaging times or higher copy numbers are required for high precision. For both single and double gradients, we demonstrate the existence of optimal length scales for the gradients, where precision is maximized, as well as analyzing how precision depends on the size of the concentration measuring apparatus. Our results provide fundamental constraints on the positional precision supplied by concentration gradients in various contexts, including both in developmental biology and also within a single cell.

## Summary

In many biological systems gradients of protein concentration provide precise positional information. Above a critical concentration a signal can be switched on by the gradient, whereas below the threshold the signal is switched off, thereby providing position dependent signalling. Such concentration gradients are, however, subject to unavoidable noise arising from intrinsic biochemical fluctuations. We therefore investigate how precisely a noisy concentration gradient can specify positional information. We analyse noisy one and two gradient models, where we find that time-averaging of concentration measurements potentially allows great precision to be achieved even with remarkably low protein copy numbers. We also find that a particular choice of the gradient decay length optimizes positional precision. Furthermore, in two-dimensional gradients, such as on a cell membrane, we find that positional precision is substantially independent of the size of the concentration measuring apparatus. We apply our results to a number of examples in both cell and developmental biology, including cell division positioning in bacteria and yeast, as well as precision gene expression in *Drosophila*.

## Introduction

To determine position in a biological system, some component within the system must have a non-uniform spatial distribution. Often this is achieved through the formation of gradients of protein concentration. Typically a gradient forms when a protein is manufactured/injected within a small region, and subsequently spreads and decays [1]. By measuring the local concentration, position relative to the source can be determined. In developmental biology, where such gradients are used to control patterns of gene expression, gradient proteins are called morphogens. However, intracellular concentration gradients are also thought to be important for organisation inside single cells.

For a gradient mechanism to be biologically viable, position determination must be precise and therefore robust to noise. Variability from one copy of the system to another (e.g. from cell to cell or embryo to embryo) will certainly compromise positional precision. Production and degradation rates can vary, for example, due to different copy numbers of transcription factors or proteases. The physical size of the system will also vary and this may affect proper positioning. Most previous analyses of morphogen gradients have focused on robustness to changes in these extrinsic factors [2-4] between different copies of the system. However, there will also be intrinsic noise affecting the gradient within a single copy of the system, for example due to the unavoidably noisy nature of the biochemical reactions involved. This dissection of the fluctuations into extrinsic or intrinsic mirrors that introduced into the analysis of stochastic gene expression [5-7]. However, here intrinsic noise alters not only the overall protein copy numbers (similar to [5]), but crucially also the spatiotemporal protein distribution. Even if all extrinsic variation could be eliminated, intrinsic biochemical

noise would still lead to a fundamental limit to the precision of position determination, in a similar way to limits on the precision of protein concentration measurement [8, 9]. In this paper, we therefore address the question of how precisely a concentration gradient can specify positional information, and calculate the limits on positional precision for a simple, but biologically relevant, gradient formation mechanism with first order reaction kinetics.

Quantitative measurements, for example on the Bicoid-Hunchback system in *Drosophila* [10], have shown that remarkable positional precision can sometimes be obtained. For this reason, understanding the fundamental limits to the precision of concentration gradients is clearly an important issue in developmental biology. Our results will be equally relevant to gradients that form within single cells, where protein copy numbers of a few thousand [11-13] will lead to large density fluctuations. The properties of intracellular protein gradients have been studied by Brown and Kholodenko [14]. Recently a number of these gradients have been observed experimentally in both prokaryotic and eukaryotic systems. The bacterial virulence factor IcsA forms a polar gradient on the cell membrane of *Shigella flexneri* [15]. MipZ in *Caulobacter* forms polar gradients to aid division site selection [11]. In *B. subtilis*, the MinCD complex also forms polar gradients in order to direct division site selection to the mid-plane of the cell [16, 17]. In *E. coli* the oscillatory dynamics of the Min proteins creates a time-averaged gradient that directs cell division placement [18-24]. Using mechanisms of this sort, division site placement in bacteria can achieve an impressive precision of  $\pm 1\%$  of the cell length [25, 26]. Cell division in eukaryotic cells is also believed to be regulated by concentration gradients. For example, in fission yeast, the protein Pom1p forms a cortical concentration gradient emanating from a cell tip, thereby restricting the cell division protein Mid1p to the cell centre [27, 28]. In eukaryotic cells, gradients of the Ran and HURP proteins aid the formation of the mitotic spindle by biasing microtubule growth towards the chromosomes [29-33]. Gradients may also play a role in the localization of Cdc42 activation, thereby permitting a coupling between cell shape and protein activation [34, 35].

Suppose that a biological system needs to identify a particular position along its length, such as the mid-plane to ensure symmetrical cell division. As concrete examples, MipZ and the MinCD complex act by displacing the essential cell division protein FtsZ from the cell membrane. Since the concentrations of MipZ/MinCD are higher near the cell poles, FtsZ accumulates near the cell centre. Below some critical threshold of MinCD or MipZ concentration, enough FtsZ will presumably accumulate to form the division apparatus. The locations where the concentration gradient crosses these thresholds mark positions within the cell. In our analysis we will simply postulate the existence of such well-defined critical thresholds, where the gradient sharply switches a downstream signal from on to off. Clearly any real gradient cannot act as such a sharp switch – in reality a certain amount of smearing is inevitable. Furthermore, there will be additional noise in the process of actually measuring the concentration due both to the binding of the gradient proteins to the receptor molecules [8, 9], and also to the downstream reactions that process this incoming signal [5-7, 36-38].

In general, the noise of the output signal of a processing network can be written as the sum of a contribution from the noise in the input signal plus a contribution from the reactions that constitute the processing network. We assume here that the detector and the processing network are ideal and do not add any noise to the gradient input signal. As a result, our calculated variation constitutes a lower bound; any real gradient signalling system will inevitably have a lower precision.

We first consider a system with a single planar morphogen source and linear degradation, thereby producing an exponentially decaying average concentration profile. While this model is very simple, it remains biologically relevant in both developmental and intracellular contexts. Gradients of Bicoid in *Drosophila* and IcsA in *Shigella* have been quantitatively measured and shown to fit this exponential decay profile on average to high accuracy [10, 15]. We then calculate the expected distribution of positions where a noisy gradient crosses a concentration threshold. With typical cellular copy numbers of a few thousand proteins, the system will be unable to identify the correct threshold position from a single measurement. In order to achieve reliable position determination the concentration must be averaged over time. We show that by averaging measurements, a biological system is able to achieve precision in position determination of a few percent of the system size even with hundreds of protein copies, a result we verify by computer simulations. Furthermore, we find that the precision of position determination is maximised when a particular choice of the gradient decay length is made. We also show how the precision depends on the detector size (i.e. the volume over which the density measurement is made). For a two dimensional gradient (e.g. on a membrane), the precision possible after a certain averaging time depends only very weakly on the detector size. We relate all these results to experimental measurements of gradients in *Shigella* and fission yeast.

We also consider the ability of gradients from two poles to identify the centre of the system, as in the MipZ and Pom1p gradients discussed above. Related designs have also been proposed for the control of *hunchback* positioning in *Drosophila* [3, 4, 39]. As before, we find that the precision of the system can be optimised by a particular choice of the decay length. However, if the threshold position is set at the system centre, time-averaging improves precision more slowly than in the single-source model. For subcellular gradients we find that a few thousand copies of the gradient proteins may therefore be required for high precision. Our results strongly constrain the possible concentrations of gradient proteins in two gradient systems.

## Results

### Single Gradient Model

We consider a protein gradient which is used to determine a particular position along the length of a cylindrical system. The system will have dimension  $d = 2$

if the gradient is restricted to the membrane, or  $d = 3$  if the gradient is in the cytoplasm. We choose the  $x$ -axis along the long axis of the system. Position in the remaining coordinates is denoted by the vector  $\mathbf{y}$ . For a membrane system, periodic boundary conditions are appropriate in the  $y$ -direction. Otherwise, zero-flux boundaries are used throughout. The system length is  $L$ , and the size of the system in the remaining directions is taken to be  $L_\perp$  (so  $L_\perp = 2\pi r$ , where  $r$  is the system radius, for the  $2d$  membrane case). A source on the  $x = 0$  plane produces proteins at rate  $J$  per unit area, which then diffuse with diffusion constant  $D$ , and are degraded uniformly at rate  $\mu$ . Neglecting fluctuations, the protein concentration  $\rho(x, \mathbf{y}, t)$  will be described by

$$\frac{\partial \rho}{\partial t} = D \nabla^2 \rho - \mu \rho + J \delta(x). \quad (1)$$

If  $L \gg \lambda = \sqrt{D/\mu}$ , the characteristic decay length of the gradient, we find that, at steady state, the density is

$$\rho(x) = \frac{J\lambda}{D} \exp(-x/\lambda). \quad (2)$$

Symmetry dictates that the average density is independent of  $\mathbf{y}$ . Gradients with the form (2) have been found to accurately fit quantitatively measured concentration profiles in both developmental [10] and subcellular [15] systems.

While we have outlined the model in terms of production and degradation, (1) could equally apply to other mechanisms in which the active protein originates in a single location, but deactivation occurs uniformly throughout the system. The same equation would therefore describe a protein which is phosphorylated by a polar-localised kinase and dephosphorylated by a uniformly distributed phosphatase, or a protein which is activated by being injected into the membrane at a pole and deactivated when it dissociates. These biochemical details do not affect the behaviour of the model.

We suppose that signalling is active where the local gradient protein concentration is above some threshold value,  $\rho_T$ , and inactive otherwise. The average concentration profile for a single gradient, (2), suggests that the system will be divided into a region  $0 \leq x < x_T$  where signalling is active, and a region  $x_T \leq x \leq L$  where signalling is not active, with  $\rho_T = \rho(x_T)$ . However, noise in the local protein concentration will cause this threshold position to fluctuate. This noise may come from intrinsic fluctuations in the diffusion, injection and decay processes, or from extrinsic factors which produce systematic changes in the boundary position when comparing one copy of the system to another. Here we consider only intrinsic biochemical fluctuations.

Protein production and degradation events are considered to be single molecule reactions with a fixed probability per unit time, and hence will be Poisson processes. We also assume that the hopping of proteins in or out of a particular region of space is governed by Poisson statistics, thereby generating a diffusive process for molecular transport. Since the system is linear, the instantaneous fluctuations in molecular number,  $n$ , within a volume  $(\Delta x)^d$  centred on the

position  $(x, \mathbf{y})$  should also obey Poisson statistics, with

$$\langle n(x)^2 \rangle - \langle n(x) \rangle^2 = \langle n(x) \rangle. \quad (3)$$

In terms of protein density, this becomes

$$\langle (\Delta \rho(x))^2 \rangle = \langle \rho(x)^2 \rangle - \langle \rho(x) \rangle^2 = \frac{\langle \rho(x) \rangle}{(\Delta x)^d}. \quad (4)$$

This relation can also be established using more elaborate field theoretic techniques (see [40]). From this expression for the variation in the density we can compute the width of the threshold position distribution by expanding about the average threshold position  $x_T$ . To leading order, this width is given by

$$w_0 = \frac{\Delta \rho(x_T)}{|\langle \rho'(x_T) \rangle|} = \sqrt{\frac{\lambda D}{J(\Delta x)^d}} \exp(x_T/2\lambda), \quad (5)$$

where  $\rho'(x_T)$  denotes the first derivative of the density at  $x = x_T$ .

Here we identify  $(\Delta x)^d$  as the size of the region in which the concentration is being measured. For subcellular gradients involved in positional information, this volume will be determined by the size of an individual receptor or protein with which the gradient protein interacts, an example being the interaction between the MinCD and FtsZ proteins in *B. subtilis*. The size of the detector,  $\Delta x$ , will then be on a molecular scale. This conclusion still holds even if the gradient proteins bind cooperatively to the “detection” protein/receptor due to the close physical proximity of the bound molecules. In contrast, however, the cellular length scale will be much larger,  $1\mu\text{m}$  or bigger.

Throughout the following analysis we will focus on subcellular gradients. However, our model can equally be applied to developmental biology, and we will consider these systems further in the *Discussion*. As concrete examples we first consider the IcsA polar gradient on the membrane of the rod-shaped bacterium *Shigella* ( $L \approx 3\mu\text{m}$ ,  $L_\perp \approx 3\mu\text{m}$ ) [15]. IcsA is exported to the outer membrane at a single pole, after which it diffuses and undergoes uniform proteolysis by the protease IcsP, thereby forming an exponential gradient exactly as in our model [15]. Outer membrane IcsA is then able to recruit actin nucleation factors. However, a critical concentration of IcsA is likely needed for actin nucleation: in this way a comet-like actin tail is generated at only one cell pole thereby generating unidirectional motility of the pathogen. A cell will typically have a few thousand copies of IcsA [12], forming a gradient with  $\lambda \approx 0.5\mu\text{m}$  [15]. We take the detector size to be  $\Delta x = 0.01\mu\text{m}$ , consistent with an interaction between IcsA and actin nucleation proteins. For diffusion on the cell membrane, we take  $D = 1\mu\text{m}^2\text{s}^{-1}$ . On the membrane of a cell of this size, there would be approximately  $LL_\perp/(\Delta x)^2 \sim 10^5$  potential detector sites, many more than the typical copy number. Even near to the source pole, detector sites will typically be unoccupied. A detector region at a distance  $x = 0.5\mu\text{m}$  from the highly-occupied pole will have average occupancy of  $\langle n \rangle \sim 10^{-1}$ . In the cytoplasm of a similarly sized bacterium, the number of potential detector sites will be

$\sim 10^6$ , again much larger than the protein copy numbers typically supported by bacteria.

Similar estimates can be made for single polar gradients in fission yeast ( $L = 10\mu\text{m}$ ,  $L_{\perp} = 6\mu\text{m}$ ), such as for Pom1p [27, 28]. Here we assume a total of 2000 protein copies (this concentration has not yet been measured but this number is plausible [28]). We also take  $D = 1\mu\text{m}^2\text{s}^{-1}$  and a decay length of  $\lambda = 2\mu\text{m}$ , parameters that are approximately consistent with the Pom1p gradient imaged by Padte et al [28]. We again assume that  $\Delta x = 0.01\mu\text{m}$  corresponding to a molecular sized detector, as would be the case if the gradient protein interacted with other membrane proteins (such as Mid1p) [27, 28]. The typical occupancy of a  $\Delta x = 0.01\mu\text{m}$  site is then  $\langle n \rangle \sim 10^{-2}$  at  $x = 2\mu\text{m}$  from the source.

As we have seen for both fission yeast and *Shigella*, average detector site occupancies that are very much less than one ensure that the threshold concentration must necessarily be less than one protein per site. Since most regions will be devoid of any copies of the protein, a single instantaneous measurement of the protein density cannot give a good estimate of the local average concentration. Additionally, multiple positions where the concentration crosses  $\rho_T$  will be observed simultaneously in such a measurement since the concentration will be above the threshold everywhere there is a protein molecule present, and below the threshold where there is no protein molecule. In order to reliably determine the average concentration profile the system must therefore integrate the measured concentration over time.

The noisy concentration profile provided by the gradient protein forms the input signal that is then time-averaged by a downstream signal processing network. In general, the mechanism for time averaging is provided by the lifetimes of the states in the processing network. For instance, in the case of gene expression, fluctuations in the occupancy of the promoter by a gene regulatory protein can be filtered by the lifetime of the mRNA transcript, provided that lifetime is much longer than the timescale of fluctuations in the promoter occupancy [7, 9]. Similarly, for subcellular gradients, as in *Shigella*, fluctuations in the gradient can be filtered by the lifetime of activated receptors/detector proteins or their downstream products. Provided this time scale is much longer than the sub-millisecond timescale of the gradient fluctuations, then good time-averaging can be achieved. Importantly, the reactions in the downstream network not only time-average the noise of the input signal, but also add further noise to the signal [5-7, 36-38]. Here, we focus exclusively on noise in the concentration gradient and do not model the downstream reactions explicitly, but simply assume they are noiseless and model them with an effective averaging time. In essence we assume that the detector and the network that process the gradient signal are ideal and do not add further noise, and are thus able to time-average the gradient signal in the best possible way. Our results thus provide a lower bound to the output noise set by the Poissonian fluctuations of the signalling molecules.

We suppose that averaging over a time-interval  $\tau$  we can take  $N_{\tau} = \tau/\tau_{ind}$  independent measurements of the concentration. In our ideal case, we then expect that the fluctuations in the concentration will decrease according to

$1/\sqrt{N_\tau}$ . Since the width varies linearly with  $\Delta\rho$  according to (5), the width will also decrease as

$$w(\tau) \sim w_0 \sqrt{\frac{\tau_{ind}}{\tau}}. \quad (6)$$

The time-scale  $\tau_{ind}$  on which independent measurements can be made is set in our ideal case solely by the reaction-diffusion dynamics of the gradient proteins, as discussed in the *Appendix*. For cellular parameter values, the typical reaction timescale,  $1/\mu$ , will be much longer than the typical timescale for diffusion between detector sites,  $(\Delta x)^2/D$ . Assuming a molecular sized detector, this latter timescale will be of order  $10^{-4}s$ , whereas effective protein lifetimes will typically be seconds or longer. The Damkohler number for the system, the ratio of the diffusive and reaction timescales, will therefore be  $Da \sim (\Delta x)^2/\lambda^2 \sim 10^{-4}$ . Since  $Da \ll 1$ , the averaging time-scale is dominated by diffusive motion. In  $d = 3$  we find  $\tau_{ind} \sim (\Delta x)^2/D$ . However, in  $d = 2$ , density correlations decay away more slowly, leading to the appearance of logarithmic corrections that are weakly dependent on the parameters  $\lambda$  and  $\Delta x$ . For long averaging times,  $\tau \gg 1/\mu$ , the width determined from time-averaged measurements will be

$$w(\tau) = k_{2d} \left[ \frac{\lambda}{\tau J} \exp(x_T/\lambda) \left( \ln \left( \frac{\lambda^2}{(\Delta x)^2} \right) + \alpha \right) \right]^{1/2} \quad (7)$$

in  $d = 2$ , and for  $d = 3$

$$w(\tau) = k_{3d} \left[ \frac{\lambda}{\tau J(\Delta x)} \exp(x_T/\lambda) \right]^{1/2}, \quad (8)$$

where  $k_{2d}$ ,  $k_{3d}$  and  $\alpha$  are constants.

As we have discussed above,  $\Delta x$  will be set by the concentration detection mechanism. However, in a subcellular context,  $\Delta x$  also sets the highest possible resolution of the system. Once  $w \approx \Delta x$  the cell cannot resolve the target position with any higher precision. Equation (7) suggests that in  $d = 2$ , precision depends only very weakly on the detector size, through the logarithmic correction factor. Reducing the detector size will increase the number of independent measurements made in a given averaging time. However, since fewer proteins will be measured by each detector over one averaging period, reducing  $\Delta x$  will therefore increase the instantaneous density fluctuations. In  $d = 2$  these two effects will largely cancel. Hence, even if we have over/underestimated the detector volume, this will have little effect on the precision of two dimensional gradients, such as IcsA in *Shigella* or Pom1p in fission yeast. In three dimensions, however,  $w$  varies as  $(\Delta x)^{-1/2}$ . Since increasing  $\Delta x$  reduces  $w$  in both  $d = 2$  and  $d = 3$ , an optimal strategy would be to choose  $\Delta x$  to match the desired precision in order to minimise the required averaging time.

Intriguingly, from equations (7) and (8) we find that there exists an optimal decay length such that precision is maximised. This result can be understood as follows: for fixed  $x_T$ , and for  $\lambda \gg x_T$ , the value of the  $|\langle \rho'(x_T) \rangle|$  tends to a constant  $J/D$ , independent of  $x_T$ . However, as  $\lambda$  increases,  $\langle \rho(x_T) \rangle$  increases



and therefore the absolute size of the fluctuations in the density also increases. Therefore, for large and increasing values of  $\lambda$ ,  $w \propto \langle \sqrt{\rho(x_T)} \rangle / \langle \rho'(x_T) \rangle$  must be increasing. Now if  $\lambda$  is small ( $\lambda \ll x_T$ ) and decreasing, when computing the width  $\propto \langle \sqrt{\rho(x_T)} \rangle / \langle \rho'(x_T) \rangle$  the presence of the square root means that the numerator decreases much more slowly than the denominator. Hence the width must again increase as  $\lambda$  is decreased for small  $\lambda$ . Combining these results for small and large  $\lambda$ , the width must have a minimum, optimum value as a function of  $\lambda$ . This occurs at  $\lambda = x_T$  in  $d = 3$ . In  $d = 2$ , the optimal decay length is given approximately by

$$\lambda \approx x_T \left( 1 - \frac{1}{\ln(x_T/(\Delta x))} \right), \quad (9)$$

where we have retained the first order logarithmic correction.

In order to examine the biological impact of equation (7) we again consider the Pom1p membrane gradient in fission yeast [27, 28], using the parameters described earlier. Simulations of this example system were performed as described in the *Methods* with on average 100 proteins in the system. Figures 1A and B show how the measured threshold position,  $\bar{x}$ , and width,  $w$ , vary with averaging time. For long averaging times the simulation data gives excellent agreement with (7), with the constants  $k_{2d} = 0.40 \pm 0.02$  and  $\alpha = 2.5 \pm 0.8$ . Figure 1C shows the  $w \sim \tau^{-1/2}$  behaviour predicted in (7), and figure 1D confirms that the width has a minimum as a function of  $\lambda$ . The simulation results are consistent with the position of the minimum predicted by (9). Figure 1E shows that the distribution of measured threshold positions is Gaussian to a good approximation.

Since the averaging timescale  $\tau_{ind}$  in a subcellular system is of order  $\sim 10^{-4}s$ , time-averaging over a period of minutes can achieve great precision even with very few copies of the gradient protein. With the parameter values given above, equation (7) predicts that the position  $x_T = 2\mu m$  can be located to within  $\pm 0.5\mu m$  within an averaging time  $\tau = 60s$  even if the system contains on average only about 20 copies of the protein.  $\pm 0.1\mu m$  precision can be achieved in the same averaging time with around 400 copies of the protein, a remarkably high level of precision for such a low concentration. *In vivo* Pom1p gradients may be formed by a few thousand protein copies, allowing for even greater precision.

However, we can see in figure 1B that for averaging times of less than about a second, the simulation results are not consistent with (7). In this regime both  $w$  and  $\bar{x}$  are equal to  $\lambda$ . As discussed above, at very short averaging times the presence of a particle at any position will cause the time-averaged concentration to be above  $\rho_T$  at that point and hence generally will generate a threshold crossing. The probability distribution of threshold measurements,  $p(x)$ , will therefore follow the probability distribution of particles. Assuming  $L \gg \lambda$  we have

$$p(x)dx = \lambda^{-1} \exp(-x/\lambda)dx. \quad (10)$$

The cell will on average estimate the threshold position to be

$$\bar{x} = \int_0^L xp(x)dx \approx \lambda, \quad (11)$$

and measurements will be distributed about this position with variance

$$w^2 = \int_0^L (x - \bar{x})^2 p(x)dx \approx \lambda^2. \quad (12)$$

The system is therefore unable to resolve the correct threshold position at these short time scales if this is different from  $\lambda$ .

Associated with the average concentration at the threshold is a length scale,  $l \sim \rho_T^{-1/d}$ , the typical distance between proteins at this position. The average time for a protein to diffuse this distance will scale as  $l^2/D$ . In two dimensions, this time is given by

$$\tau_x \sim (\langle \rho(x_T) \rangle D)^{-1} = (J\lambda)^{-1} \exp(x_T/\lambda). \quad (13)$$

Since  $\tau_x$  is the timescale on which a diffusing particle first arrives at  $x_T$ , if  $\tau \ll \tau_x$  there will generally be no particles detected at  $x_T$  in the averaging period. The system therefore cannot reliably estimate the mean concentration at  $x_T$ , and hence cannot precisely identify the threshold position. For averaging times much greater than  $\tau_x$ , on average at least one particle will be detected at  $x_T$ . The time-averaged concentration profile will then approach (2), and  $\bar{x}$  will approach  $x_T$ . Hence  $\tau_x$  determines the cross-over time between the two observed regimes of constant  $w$  and  $w \propto \tau^{-1/2}$ . Figure 1F shows that the scaling in equation (13) is also reproduced in our simulations. For the parameter values above,  $\tau_x = 0.3s$ , and for a more realistic copy number of 1000,  $\tau_x = 0.03s$ . These timescales are extremely short compared to cell cycle timescales, but do nevertheless show that some sort of time averaging is probably essential: a single instantaneous measurement is unlikely to provide precise positional information. In fact, as we have seen, averaging over much longer times (tens of seconds) may be necessary if very high (1%) precision is required.

Simulations of the model in three dimensions were also performed (data not shown). Similar behaviour was observed in this case, and equation (8) gave good agreement with the observed width at long averaging times.

## Oppositely directed gradients

In order to reliably locate the centre of a system, the mechanism responsible must incorporate information about the overall system size so that the identified position can scale correctly. A single gradient characterised by a fixed decay length cannot achieve this. We therefore examine a system where protein gradients are produced by sources at both ends, and where the central position is identified as a concentration minimum.

We modify our earlier model by adding an additional planar source at  $x = L$ . This addition is appropriate for modelling cell division inhibitors, such as MipZ

in *Caulobacter*, that are injected into the membrane near both cell poles. However, our model would apply equally if the two sources are of different repressor proteins (as may be the case in fission yeast [27, 28]), although we do assume that  $J$ ,  $D$  and  $\mu$  are the same for both gradients. In this scenario, signalling activity will be determined by the total concentration. Without fluctuations, this will be described by

$$\frac{\partial \rho}{\partial t} = D \nabla^2 \rho - \mu \rho + J \delta(x) + J \delta(x - L). \quad (14)$$

The steady-state solution is now

$$\rho(x) = \frac{J\lambda}{D} \frac{\cosh((x - L/2)/\lambda)}{\sinh(L/2\lambda)}, \quad (15)$$

which has the expected minimum at  $x = L/2$ .

We then suppose that the cell compares the concentration to a threshold value corresponding to the minimum of the average profile,  $\rho_{min} = \rho(L/2) = \rho_T$ . Positions where the concentration is at or below the threshold are identified as being at the centre of the cell. While the average steady-state density profile would never extend below  $\rho_{min}$ , fluctuations ensure that the concentration in the region around the centre spends a significant amount of time at or below the threshold. Around point(s) where  $\langle \rho(x) \rangle = \rho_T$ , noise in the protein concentration will lead to a distribution of threshold crossing positions. We consider an expansion of the density fluctuations about  $x_T = L/2$ , giving, to leading order

$$\Delta \rho(x_T) = \frac{1}{2} |\langle \rho''(x_T) \rangle| w^2, \quad (16)$$

since any first order term proportional to  $\langle \rho' \rangle$  vanishes at  $x_T = L/2$ . The width is therefore given by

$$w^2 = \frac{2\Delta \rho(L/2)}{\langle \rho''(L/2) \rangle}. \quad (17)$$

Substituting in (15) gives

$$w_0 = \left( \frac{4D\lambda^3 \sinh(L/2\lambda)}{J(\Delta x)^d} \right)^{1/4}. \quad (18)$$

As in the single gradient model, the typical occupancy of the threshold region will be much less than one. For example, if we take the parameter values considered previously for the Pom1p gradient in fission yeast, with 2000 protein copies, the average occupancy of a detector site at  $x = L/2$  will be  $\langle n(L/2) \rangle \sim 10^{-3}$ . We assume here that Pom1p forms a gradient from both poles. In fact it may only form a single gradient with another hitherto unidentified protein forming the second polar gradient [27, 28]. However, as discussed earlier, this detail does not affect our calculations. As a second example, MipZ in *Caulobacter* ( $L = 2.5\mu\text{m}$ ,  $L_\perp = 2\mu\text{m}$ ) is typically present at about 1000 copies, and forms two polar gradients with a decay length  $\lambda \approx 0.25\mu\text{m}$  [11]. The average

occupancy at the centre of this system would be approximately  $\langle n(L/2) \rangle \sim 10^{-3}$ . Averaging measurements of the concentration over time is therefore required in both cases to obtain precise positional information. Since the width now goes as  $(\Delta\rho)^{1/2}$ , as shown in (17), we expect

$$\begin{aligned} w(\tau) &= w_0 \left( \frac{\tau_{ind}}{\tau} \right)^{1/4} \\ &= \begin{cases} \tilde{k}_{2d} \left[ \frac{\lambda^3}{\tau J} \sinh(L/2\lambda) \left( \ln \left( \frac{\lambda^2}{(\Delta x)^2} \right) + \tilde{\alpha} \right) \right]^{1/4} & \text{in } d = 2 \\ \tilde{k}_{3d} \left[ \frac{\lambda^3}{\tau J(\Delta x)} \sinh(L/2\lambda) \right]^{1/4} & \text{in } d = 3 \end{cases}, \end{aligned} \quad (19)$$

where  $\tilde{k}_{2d}$ ,  $\tilde{\alpha}$  and  $\tilde{k}_{3d}$  are constants. Averaging proceeds much more slowly than previously, with a  $\tau^{-1/4}$  dependence. This follows directly from the vanishing of the first derivative at the average threshold position. In  $d = 3$ , and for  $\lambda \ll L$ , equation (19) predicts that  $w$  will be minimised when  $\lambda \approx L/6$  is chosen. In  $d = 2$  logarithmic corrections again alter this result slightly, with the optimal decay length now occurring at

$$\lambda \approx \frac{L}{6} \left( 1 - \frac{1}{3 \ln(L/6(\Delta x))} \right), \quad (20)$$

where we have included the leading logarithmic correction. This result arises for similar reasons as in the single gradient model. For the Pom1p gradient imaged by Padte et al [28], the decay length is observed to be  $1 - 1.5\mu\text{m}$ , comparable to this optimal decay length of about  $1.5\mu\text{m}$  for a  $10\mu\text{m}$  cell.

We simulated our model in two dimensions with representative parameter values for fission yeast membrane gradients. We used  $\mu = 0.36\text{s}^{-1}$  chosen to give  $\lambda = 1.67\mu\text{m}$ , and  $J = 6\mu\text{m}^{-1}\text{s}^{-1}$  giving on average 200 protein copies in total. Figure 2 shows the results of these simulations. Again we observe two distinct regimes. At averaging times longer than about a second, there is excellent agreement with equation (19), as we can see in figure 2C. Fitting to the simulation results we find  $\tilde{k}_{2d} = 0.63 \pm 0.02$  and  $\tilde{\alpha} = 2.5 \pm 1.0$ . Figure 2D confirms the existence of the optimal decay length in our simulations.

Since the width decays as  $\tau^{-1/4}$  for this system, longer averaging times and/or higher protein copy numbers are required than in the single gradient model to achieve high precision. Intrinsic biochemical noise may therefore strongly constrain systems of this type. In order for the yeast-membrane gradient considered above to achieve precision of  $\pm 5\%$  of the cell length after averaging for one minute, about 800 protein copies are required. Therefore, in the absence of any other positioning mechanisms, the Pom1p gradient will require  $\sim 1000$  protein copies or more to precisely direct the location of cell division. We estimate that the MipZ gradient in *Caulobacter*, with 1000 protein copies, would be able to locate the cell centre to within  $\pm 5\%$  of  $L$  after approximately  $\tau = 2\text{s}$ . However, since precision only improves as  $\tau^{-1/4}$ , averaging over  $\tau = 20$  minutes would be required for the same system to achieve  $\pm 1\%$  accuracy.

## Discussion

Noise in biochemical processes within a cell will lead to fluctuations in protein concentration gradients, and hence also to variation in the position where these gradients cross a particular threshold value. These fluctuations therefore place a limit on the potential precision of position determination mechanisms relying on concentration gradients alone. In subcellular systems with protein copy numbers in the thousands, this noise will be sufficiently large that position cannot be determined reliably from a single measurement of the density profile. In order to determine position to within a few percent, a precision achieved by some subcellular systems, the protein concentration must be averaged over time. For a single subcellular membrane gradient, we have seen that by averaging over a period of a minute, excellent precision can potentially be achieved with only a few hundred protein copies. This remarkable precision is due to the sub-millisecond diffusive time-scale on which time-averaging occurs. Precise identification of the cell mid-plane by gradients emanating from both poles requires longer averaging times or higher copy numbers, since larger fluctuations result from the vanishing first derivative of the average concentration at the system centre. Intrinsic biochemical noise may therefore be a strong constraint on subcellular two-gradient positioning systems, dictating that the copy numbers be sufficiently high to suppress fluctuations.

So far we have focused almost exclusively on fluctuations in subcellular gradients, however our results are also applicable to developmental biology and we wish to briefly comment on this application. Here the appropriate length scales are usually much longer, on the order of hundreds of micrometers in *Drosophila*. Moreover, the gradients affect patterns of gene expression through the binding of gradient molecules to DNA regulatory sequences inside individual nuclei. For example, in *Drosophila*, where exponential gradients have been quantitatively measured for Bicoid [10], Bicoid binds cooperatively to *hunchback* regulatory DNA. In this case we again expect molecular-scale effective measuring volumes, with  $\Delta x \sim 0.01\mu\text{m}$  being a reasonable order of magnitude. We next assume purely Poisson statistics for the fluctuations: this is a stronger assumption than for our earlier subcellular gradients, as there will be additional complications arising, for example, from the import/export of morphogens from nuclear compartments. However, if diffusive noise is dominant then Poisson statistics will be retained and we can expect our earlier analysis to apply, although with one important distinction. Instead of  $\Delta x$  setting the maximal possible precision, this will now be set by the size of individual nuclei (prior to cellularization), since we expect relatively homogeneous gene expression within a single nuclear volume. A single nucleus in *Drosophila* has a length scale of around  $10\mu\text{m}$ , still much smaller than the decay length of the gradient of  $\lambda \sim 100\mu\text{m}$ , allowing for high precision gene expression [10]. Using the *Drosophila* Bicoid gradient as an example, we use  $L = 500\mu\text{m}$ ,  $L_{\perp} = 100\mu\text{m}$ , and estimate  $D = 10\mu\text{m}^2\text{s}^{-1}$  and  $\mu = 10^{-3}\text{s}^{-1}$ , giving  $\lambda = 100\mu\text{m}$ , consistent with experiment [10]. Assuming a high copy number of  $10^7$  per embryo (we are not aware of experimental constraints on this figure), gives  $J \sim 1\mu\text{m}^{-2}\text{s}^{-1}$ . For a single gradient in three

dimensions, we find that about a 5 minute averaging time is required to bring the error down to plus or minus a single nuclear length. For a two gradient model in three dimensions, longer averaging times on the order of an hour are required to reduce the centre-finding positional error to plus or minus about 2 nuclear lengths. Since gene expression may need to be controlled on shorter timescales than this, other designs, for example using *interacting* gradients [3, 4], may be required for high precision centre finding (see also below). The effects of the optimum gradient length scale will also be interesting to probe in a developmental biology context. However, our simple analysis may be complicated by the multiple roles played by many morphogens: for example, Bicoid not only activates *hunchback*, but it also helps to regulate pair-rule genes, such as Even-skipped. Nevertheless, it is interesting to note that the Bicoid gradient length scale  $\lambda \sim 100\mu\text{m}$  [10] is not too far away from the  $L/6$  optimum for a two gradient case, and in a single gradient context will offer maximal precision well into the anterior half of the embryo.

Up to this point we have only considered systems with first order degradation. Morphogen gradients with nonlinear decay have also been proposed [2]. This nonlinearity will lead to non-Poissonian density fluctuations, which may significantly change the observed behaviour. England and Cardy [41] have previously calculated the response of a gradient with nonlinear decay to one source of biochemical noise, namely a fluctuating production rate. However, they calculated the change to the average gradient, while fluctuations about this average may also be important. It would certainly be of interest to compare the performance of linear and nonlinear degradation mechanisms in more detail. Centre-finding mechanisms with interactions have also been proposed [3, 4]. In these models position is determined from the combined gradient of two proteins, which will be steep around the system centre due to an interaction between the two gradients. These mechanisms may therefore be able to achieve greater precision for mid-point determination than the noninteracting mechanism considered here.

Throughout this work we have assumed that the gradient protein concentration fluctuates about a steady-state profile, and hence averaging over a longer time will give a more precise estimate of the average profile. For a subcellular system, the steady-state gradient will develop over timescales of less than about a minute, due to the micrometer length scales involved. This timescale is short compared to the cell cycle time, which ranges from tens of minutes up to many hours. For this reason we expect that subcellular gradients will be in steady-state and therefore that our analysis will be directly applicable. However, in developmental biology, the effective lifetimes will likely be much longer, and the gradient may take hours to fully reach steady-state. Moreover, a number of developmental biology systems are known to respond to a morphogen gradient that has not reached steady-state [42-44]. A further complication is the possibility of gradient formation by non-Fickian diffusion [45], where there is no steady-state at all. The model considered in this paper does not take into account time-varying average gradients. If the average gradient is evolving, a longer averaging period will not necessarily lead to improved precision. Clearly,

more work will be required to understand how such dynamically evolving systems are able to yield precise positional information and filter out fluctuations. Nevertheless, we do note that two gradient systems of the kind analyzed here are naturally able to locate the system centre even without being in steady-state, due to the symmetry of the system [3]. The positional variations in such a non-steady-state scenario will not be the same as calculated here, but our analysis does form a first step towards the analysis of these more complex systems.

## Methods

### Calculation of $\tau_{ind}$

We have assumed in our analysis that during the time-averaging process we are taking independent measurements at intervals of  $\tau_{ind}$ . However, in both real biological systems and in our simulations, measurements can generally be taken at much shorter intervals than this, leading to correlations between consecutive measurements. For a series of correlated measurements taken at time intervals  $\delta t$  over a period  $0 \leq t \leq \tau$ , with  $\tau \gg \delta t$ , the expected error for the time-averaged concentration at position  $\mathbf{x}$ ,  $(\Delta\rho(\mathbf{x}, \tau))^2$ , is given by [46]

$$(\Delta\rho(\mathbf{x}, \tau))^2 = \frac{\delta t}{\tau} (\Delta\rho(\mathbf{x}, 0))^2 \left[ 1 + \frac{2}{\delta t} \int_0^\tau \left( 1 - \frac{t}{\tau} \right) C(t) dt \right], \quad (21)$$

where  $(\Delta\rho(\mathbf{x}, 0))^2$  is the variance of a single measurement,

$$(\Delta\rho(\mathbf{x}, 0))^2 = \langle \rho(\mathbf{x}, 0)^2 \rangle - \langle \rho(\mathbf{x}, 0) \rangle^2, \quad (22)$$

and  $C(t)$  is the normalized density correlation function,

$$C(t) = \frac{\langle \rho(\mathbf{x}, t) \rho(\mathbf{x}, 0) \rangle - \langle \rho(\mathbf{x}, 0) \rangle^2}{\langle \rho(\mathbf{x}, 0)^2 \rangle - \langle \rho(\mathbf{x}, 0) \rangle^2}. \quad (23)$$

We therefore define the timescale  $\tau_{ind}$  to be

$$\tau_{ind}(\tau) = 2 \int_0^\tau \left( 1 - \frac{t}{\tau} \right) C(t) dt, \quad (24)$$

and assuming  $\tau_{ind} \gg \delta t$  we recover

$$\Delta\rho(\mathbf{x}, \tau) = \Delta\rho(\mathbf{x}, 0) \left( \frac{\tau_{ind}(\tau)}{\tau} \right)^{1/2}. \quad (25)$$

For  $N$  independent measurements of the density, we would expect the error to decline as  $N^{-1/2}$ . For large enough values of  $\tau_{ind}(\tau)$ , where  $\tau_{ind}$  becomes independent of  $\tau$ , we can therefore interpret  $\tau_{ind}$  as the time-interval required for successive measurements to be independent.

The next step of the calculation is to compute the correlation function  $C(t)$  appropriate for our model. For pure diffusion, we expect:

$$C(t) \sim 1 \quad \text{for } t \ll \frac{(\Delta x)^2}{D} \quad (26)$$

$$C(t) \sim \left( \frac{(\Delta x)^2}{Dt} \right)^{d/2} \quad \text{for } t \gg \frac{(\Delta x)^2}{D}. \quad (27)$$

On time scales  $t \ll (\Delta x)^2/D$  the system remains perfectly correlated as there has been insufficient time for particles to hop away to neighboring sites. However, for  $t \gg (\Delta x)^2/D$ , an algebraically decaying correlation function is found, characteristic of diffusion. However, we also need to incorporate the effects of spontaneous decay that occur independently of the diffusive motion. Adding decay to the system simply alters the correlation functions by a multiplicative factor of  $\exp(-\mu t)$ . We now substitute this full form into the definition of  $\tau_{ind}$  (24). In the biologically relevant limits where  $\tau \gg (\Delta x)^2/D$  and  $1/\mu \gg (\Delta x)^2/D$ , we find, for  $d = 2$

$$\tau_{ind} \sim \frac{(\Delta x)^2}{D} \left( \ln \left( \frac{D\tau}{(\Delta x)^2} \right) + \text{constant} \right) \quad (28)$$

for  $\mu\tau \ll 1$ , and

$$\tau_{ind} \sim \frac{(\Delta x)^2}{D} \left( \ln \left( \frac{\lambda^2}{(\Delta x)^2} \right) + \text{constant} \right) \quad (29)$$

for  $\mu\tau \gg 1$ . In a three-dimensional system we find

$$\tau_{ind} \sim \frac{(\Delta x)^2}{D}. \quad (30)$$

For the parameter values considered in our simulations we do not observe the logarithmic  $\tau$ -dependence in the width predicted by (28). In the single gradient simulations this is because, at short times  $\tau \ll \tau_x$ , we enter the constant  $w \sim \lambda$  regime. For the parameter values used, the transition from  $w \sim \lambda$  at  $\tau \ll \tau_x \approx 0.3s$  to the long time behaviour (7) for  $\tau \gg 1/\mu \approx 4s$  overwhelms the small logarithmic effect. If the production rate  $J$  were increased significantly,  $\tau_x \propto J^{-1}$  would be reduced and the  $\ln \tau$  regime would become accessible since the  $\tau_x$  and  $1/\mu$  timescales would then become better separated. However, even in this case, the logarithmic variation in (28) is intrinsically weak, and will likely have a negligible effect in a biological context.

## Simulations

Stochastic simulations were performed on a two-dimensional square lattice with  $N_x = L/\delta x$  sites in the  $x$ -direction and  $N_y = L_\perp/\delta x$  sites in the  $y$ -direction, where  $\delta x = 0.01\mu\text{m}$  is the lattice spacing. The detector size  $\Delta x$  was normally set equal to  $\delta x$  except for cases where the detector size was varied, in which



case  $\Delta x$  was set to be a multiple of  $\delta x$ . Zero-flux boundaries were implemented at  $x = 0$  and  $x = L$ , and a periodic boundary was used to connect  $y = 0$  with  $y = L_{\perp}$ . A fixed time step,  $\delta t = 2.5 \times 10^{-5} s$ , was chosen so that for the given diffusion constant the total probability of diffusion out of a site in all directions approached 1. However, a timestep 5 times smaller was also tested with no effect on any of the results. For each  $x = 0$  site, particles were injected at each time step in a Poisson process with mean  $j = J\delta x\delta t$ . In the two-gradient model, particles were also added at  $x = L$  in an identical but uncorrelated process. Diffusion and decay were also treated as Poisson processes, with hopping and decay probabilities of  $D\delta t/(\delta x)^2$  and  $\mu\delta t$  per particle respectively. Simulations were initialised with the mean number of particles in the system,  $JL_{\perp}/\mu$  for the one-gradient model or twice this value for the two-gradient model, with a probability distribution that followed the average density distribution.

The mean occupancy for each detector site was calculated over the averaging period,  $\tau$ . For each site this mean occupancy was compared with each neighbouring site. If one occupancy was above the threshold and the other below, this boundary was identified as a threshold crossing position. This process was repeated for many averaging periods, ranging from  $10^5$  repeats for short averaging times to 500 repeats for very long averaging times, to generate a distribution of crossing positions throughout the system. Threshold crossings in both the  $x$ - and  $y$ -directions were observed. We found that the distributions as a function of  $x$ -position of these two types of crossing were the same. For each row of sites,  $x = 0$  to  $x = L$  at fixed  $y$ , the mean (“measured threshold”) and root-mean-squared deviation (“width”) of the threshold distribution from many averaging periods were calculated independently. In the figures we plot the mean of these two quantities across the different  $y$ -values within the system, with error bars of one standard deviation.

For the single-source model the standard parameter values used in the simulations were as follows:  $L = 10\mu m$ ,  $L_{\perp} = 6\mu m$ ,  $D = 1\mu m^2 s^{-1}$ ,  $\mu = 0.25 s^{-1}$ ,  $J = 4.17\mu m^{-1} s^{-1}$ ,  $\Delta x = 0.01\mu m$ ,  $x_T = 2\mu m$ . To generate the data collapse in figures 1C and F, simulations were also performed with:  $D = 0.5\mu m^2 s^{-1}$ ;  $J = 6.25\mu m^{-1} s^{-1}$ ;  $\Delta x = 0.02\mu m$ ;  $\mu = 1 s^{-1}$ ;  $\mu = 0.11 s^{-1}$ ;  $x_T = 1\mu m$ ;  $x_T = 3\mu m$ . For the two-source model, standard parameters were the same as above except  $\mu = 0.36 s^{-1}$  and  $J = 6\mu m^{-1} s^{-1}$ . In figure 2C data are also shown with:  $D = 0.5\mu m^2 s^{-1}$ ;  $\mu = 1 s^{-1}$ ;  $\mu = 0.25 s^{-1}$ ;  $J = 9\mu m^{-1} s^{-1}$ ;  $\Delta x = 0.02\mu m$ ;  $L = 7.5\mu m$ ;  $L = 15\mu m$  and  $\Delta x = 0.02\mu m$ .

## Funding

F.T. is supported by the EPSRC, and M.H. by The Royal Society.

## References

- [1] Wolpert L (1969) Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* 25:1–47.
- [2] Eldar A, Rosin D, Shilo BZ, Barkai N (2003) Self-Enhanced Ligand Degradation Underlies Robustness of Morphogen Gradients. *Dev Cell* 5:635–646.
- [3] Howard M, ten Wolde PR (2005) Finding the centre reliably: robust patterns of developmental gene expression. *Phys Rev Lett* 95:208103.
- [4] McHale P, Rappel WJ, Levine H (2006) Embryonic pattern scaling achieved by oppositely directed morphogen gradients. *Phys Biol* 3:107–120.
- [5] Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic Gene Expression in a Single Cell. *Science* 297:1183–1186.
- [6] Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci USA* 98:8614–8619.
- [7] Paulsson J (2004) Summing up the noise in gene networks. *Nature* 427:415–418.
- [8] Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys J* 20:193–217.
- [9] Bialek W, Setayeshgar S (2005) Physical limits to biochemical signalling. *Proc Natl Acad Sci USA* 102:10040–10045.
- [10] Houchmandzadeh B, Wieschaus E, Leibler S (2002) Establishment of developmental precision and proportions in the early *Drosophila* embryo. *Nature* 415:798–802.
- [11] Thanbichler M, Shapiro L (2006) MipZ, a Spatial Regulator Coordinating Chromosome Segregation with Cell Division in *Caulobacter*. *Cell* 126:147–162.
- [12] Magdalena J, Goldberg MB (2002) Quantification of *Shigella* IcsA Required for Bacterial Actin Polymerization. *Cell Motil Cytoskeleton* 51:187–196.
- [13] Shih YL, Fu X, King GF, Le T, Rothfield LI (2002) Division site placement in *E.coli*: mutations that prevent formation of the MinE ring lead to loss of the normal midcell arrest of growth of polar MinD membrane domains. *EMBO J* 21:3347–3357.
- [14] Brown GC, Kholodenko BN (1999) Spatial gradients of cellular phosphoproteins. *FEBS Lett* 457:452–454.
- [15] Robbins JR, Monack D, McCallum SJ, Vegas A, Pham E, et al. (2001) The making of a gradient: IcsA (VirG) polarity in *Shigella flexneri*. *Mol Microbiol* 41:861–872.

- [16] Marston AL, Errington J (1999) Selection of the midcell division site in *Bacillus subtilis* through MinD-dependent polar localization and activation of MinC. *Mol Microbiol* 33:84–96.
- [17] Howard M (2004) A Mechanism for Polar Protein Localization in Bacteria. *J Mol Biol* 335:655–663.
- [18] Raskin DM, de Boer PAJ (1999) Rapid pole-to-pole oscillation of a protein required for directing division to the middle of *Escherichia coli*. *Proc Natl Acad Sci USA* 96:4971–4976.
- [19] Meinhardt M, de Boer PAJ (2001) Pattern formation in *Escherichia coli*: A model for the pole-to-pole oscillations of min proteins and the localization of the division site. *Proc Natl Acad Sci USA* 98:14202–14207.
- [20] Howard M, Rutenberg AD, de Vet S (2001) Dynamic Compartmentalization of Bacteria: Accurate Division in *E. Coli*. *Phys Rev Lett* 87:278102.
- [21] Kruse K (2002) A Dynamic Model for Determining the Middle of *Escherichia coli*. *Biophys J* 82:618–627.
- [22] Huang KC, Meir Y, Wingreen NS (2003) Dynamic structures in *Escherichia coli*: Spontaneous formation of MinE rings and MinD polar zones. *Proc Natl Acad Sci USA* 100:12724–12728.
- [23] Howard M, Kruse K (2005) Cellular organization by self-organization: mechanisms and models for Min protein dynamics. *J Cell Biol* 168:533–536.
- [24] Tostevin F, Howard M (2006) A stochastic model of Min oscillations in *Escherichia coli* and Min protein segregation during cell division. *Phys Biol* 3:1–12.
- [25] Yu XC, Margolin W (1999) FtsZ ring clusters in *min* and partition mutants: role of both the Min system and the nucleoid in regulating FtsZ ring localization. *Mol Microbiol* 32:315–326.
- [26] Migocki MD, Freeman MK, Wake RG, Harry EJ (2002) The Min system is not required for precise placement of the midcell Z ring in *Bacillus subtilis*. *EMBO Rep* 3:1163–1167.
- [27] Celton-Morizur S, Racine V, Sibarita JB, Paoletti A (2006) Pom1 kinase links division plane position to cell polarity by regulating Mid1p cortical distribution. *J Cell Sci* 119:4710–4718.
- [28] Padte NN, Martin SG, Howard M, Chang F (2006) The Cell-End Factor Pom1p Inhibits Mid1p in Specification of the Cell Division Plane in Fission Yeast. *Curr Biol* 16:2480–2487.
- [29] Kalb P, Weis K, Heald R (2002) Visualization of a Ran-GTP Gradient in Interphase and Mitotic *Xenopus* Egg Extracts. *Science* 295:2452–2456.

- [30] Niethammer P, Bastiaens P, Karsenti E (2004) Stathmin-Tubulin Interaction Gradients in Motile and Mitotic Cells. *Science* 303:1862–1866.
- [31] Caudron M, Bunt G, Bastiaens P, Karsenti E (2005) Spatial Coordination of Spindle Assembly by Chromosome-Mediated Signaling Gradients. *Science* 309:1373–1376.
- [32] Wollman R, Cytrynbaum EN, Jones JT, Meyer T, Scholey JM, et al. (2005) Efficient Chromosome Capture Requires a Bias in the ‘Search-and-Capture’ Process during Mitotic-Spindle Assembly. *Curr Biol* 15:828–832.
- [33] Wong J, Fang G (2006) HURP controls spindle dynamics to promote proper interkinetochore tension and efficient kinetochore capture. *J Cell Biol* 173:879–891.
- [34] Meyers J, Craig J, Odde DJ (2006) Potential for Control of Signaling Pathways via Cell Size and Shape. *Curr Biol* 16:1685–1693.
- [35] Howard M (2006) Cell Signalling: Changing Shape Changes the Signal. *Curr Biol* 16:R673–R675.
- [36] Detwiler PB, Ramanathan S, Sengupta A, Shraiman BI (2000) Engineering Aspects of Enzymatic Signal Transduction: Photoreceptors in the Retina. *Biophys J* 79:2801–2817.
- [37] Shibata T, Fujimoto K (2005) Noisy signal amplification in ultrasensitive signal transduction. *Proc Natl Acad Sci USA* 102:331–336.
- [38] Tanase-Nicola S, Warren PB, ten Wolde PR (2006) Signal Detection, Modularity, and the Correlation between Extrinsic and Intrinsic Noise in Biochemical Networks. *Phys Rev Lett* 97:068102.
- [39] Houchmandzadeh B, Wieschaus E, Leibler S (2005) Precise domain specification in the developing *Drosophila* embryo. *Phys Rev E* 72:061920.
- [40] Täuber UC, Howard M, Vollmayr-Lee BP (2005) Applications of field-theoretic renormalization group methods to reaction-diffusion problems. *J Phys A: Math Gen* 38:R79–R131.
- [41] England JL, Cardy J (2005) Morphogen Gradient from a Noisy Source. *Phys Rev Lett* 94:078101.
- [42] Harfe BD, Scherz PJ, Nissim S, Tian H, McMahon AP, et al. (2004) Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* 118:517–258.
- [43] Ahn S, Joyner AL (2004) Dynamic changes in the response of cells to positive hedgehog signalling during mouse limb partitioning. *Cell* 118:505–516.

- [44] Park HC, Shin J, Appel B (2004) Spatial and temporal regulation of ventral spinal cord precursor specification by Hedgehog signalling. *Development* 131:5959–5969.
- [45] Horning G, Berkowitz B, Barkai N (2005) Morphogen gradient formation in a complex environment: An anomalous diffusion model. *Phys Rev E* 72:041916.
- [46] Müller-Krumbhaar H, Binder K (1973) Dynamic Properties of the Monte Carlo Method in Statistical Mechanics. *J Stat Phys* 8:1–24.

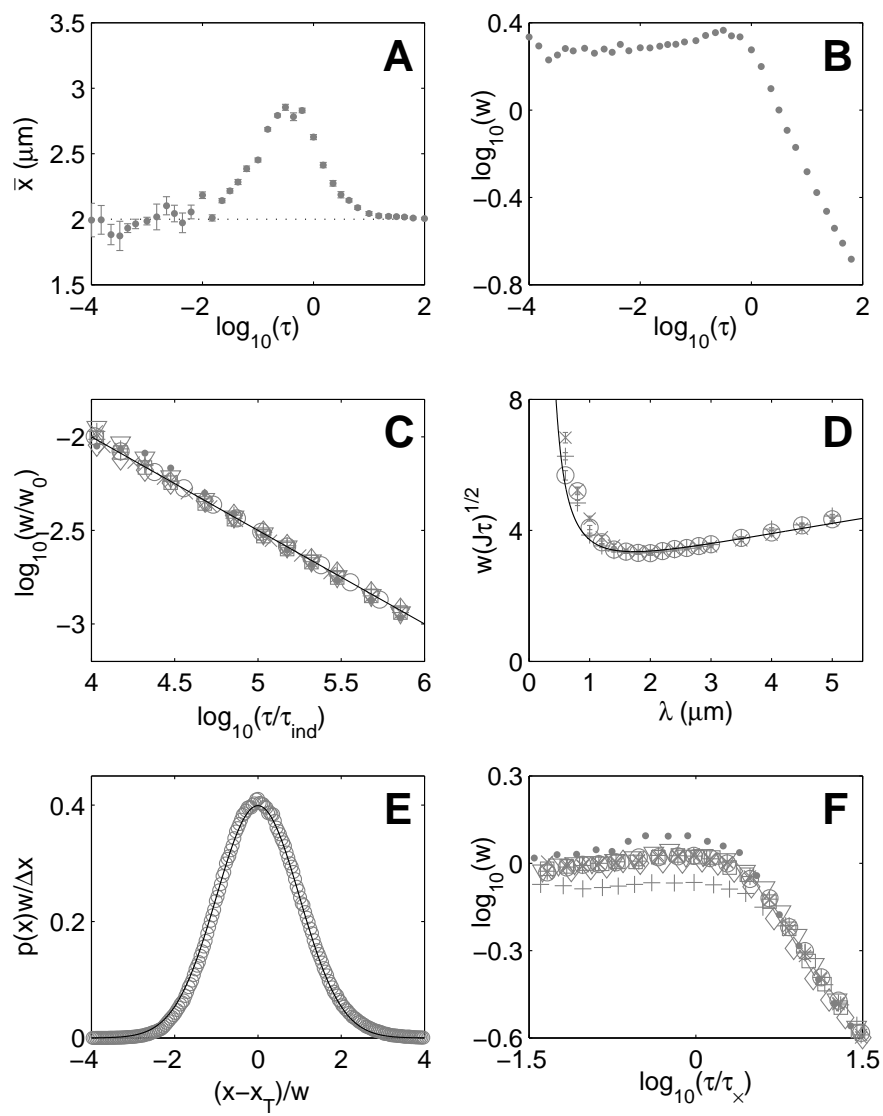


Figure 1: Single gradient model in 2d. A. Variation of the estimated threshold position with averaging time, with  $x_T = 2\mu\text{m}$  and  $\lambda = 2\mu\text{m}$ . B. Variation of the width as a function of averaging time. C. Data collapse of the width at large  $\tau$  for a range of parameter values. Full line shows the prediction of equation (7) with  $k_{2d} = 0.40$  and  $\alpha = 2.5$ . D.  $w(\tau)$  as a function of decay length, with  $x_T = 2\mu\text{m}$ . Results for three different averaging times are shown:  $\times$ :  $\tau = 10\text{s}$ ;  $\circ$ :  $\tau = 15\text{s}$ ; and  $+$ :  $\tau = 22.5\text{s}$ . The full line shows the prediction from equation (7). At large  $\lambda$  the simulation results deviate from the prediction since the assumption that  $L \gg \lambda$  is no longer valid. E. Plot of the probability distribution for measuring the threshold at position  $x$  with an averaging time  $\tau = 45\text{s}$ . The full line shows a normal distribution. F. Scaling of the cross-over time,  $\tau_\times$ , according to equation (13). In figures A., B. and E. the standard parameter values given in the text were used. In figures C. and F., \* indicates the standard parameter values. For the other data sets one parameter value was changed as follows:  $\circ$ :  $D = 0.5\mu\text{m}^2\text{s}^{-1}$ ;  $\square$ :  $J = 6.25\mu\text{m}^{-1}\text{s}^{-1}$ ;  $\times$ :  $\Delta x = 0.02\mu\text{m}$ ;  $\bullet$ :  $\mu = 1\text{s}^{-1}$ ;  $+$ :  $\mu = 0.11\text{s}^{-1}$ ;  $\diamond$ :  $x_T = 1\mu\text{m}$ ;  $\nabla$ :  $x_T = 3\mu\text{m}$ .

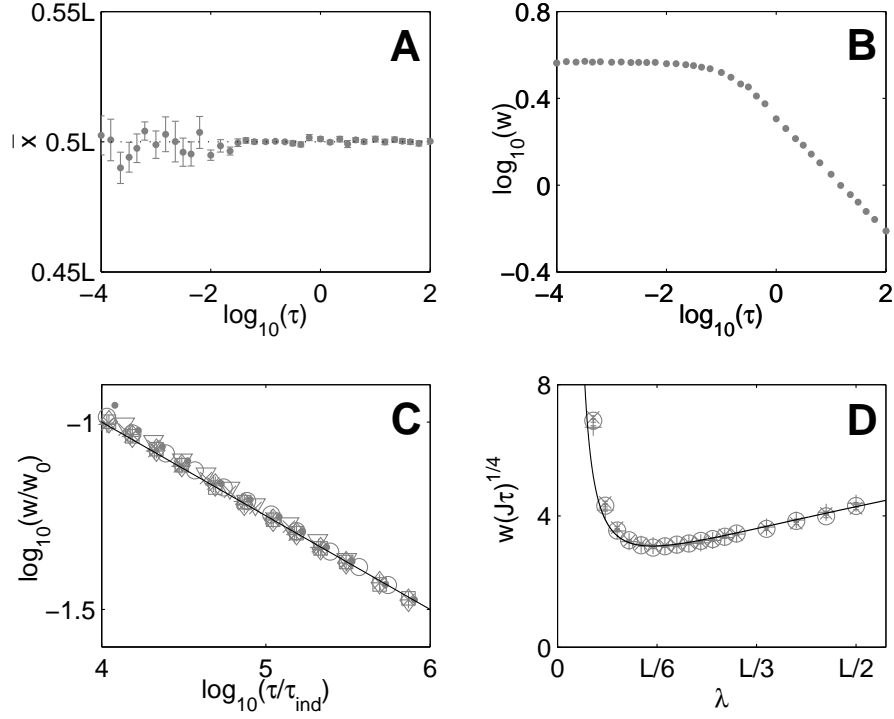


Figure 2: Two gradient model in 2d. A. The mean threshold position fluctuates about  $L/2$  due to the symmetry of the system. B. Variation of the width  $w$  as a function of averaging time. C. Data collapse of the width as a function of averaging time, at long times, for a range of parameter values. The full line shows (19) with  $k_{2d} = 0.63$  and  $\tilde{\alpha} = 2.5$ . \* indicates the standard parameter values. For the other data sets parameter values were changed as follows:  $\circ$ :  $D = 0.5\mu\text{m}^2\text{s}^{-1}$ ;  $\square$ :  $J = 9\mu\text{m}^{-1}\text{s}^{-1}$ ;  $\times$ :  $\Delta x = 0.02\mu\text{m}$ ;  $\bullet$ :  $\mu = 1\text{s}^{-1}$ ;  $+$ :  $\mu = 0.25\text{s}^{-1}$ ;  $\diamond$ :  $L = 7.5\mu\text{m}$ ;  $\nabla$ :  $L = 15\mu\text{m}$  and  $\Delta x = 0.02\mu\text{m}$ . D. Plot of width as a function of decay length for averaging times  $\times$ :  $\tau = 30\text{s}$ ;  $\circ$ :  $\tau = 45\text{s}$ ; and  $+$ :  $\tau = 60\text{s}$ . The full line shows the prediction from equation (19).